MAGIC: An Experimental System for Generating Multimedia Briefings about Post-Bypass Patient Status

Mukesh Dalal, Ph.D., Steven Feiner, Ph.D., Kathleen McKeown, Ph.D.
Department of Computer Science
Columbia University
New York, NY 10027

Desmond Jordan, M.D., Barry Allen, Ph.D.
Anesthesiology and Medical Informatics Depts.
College of Physicians and Surgeons
Columbia University, New York, NY 10032

Yasser alSafadi, Ph.D.
Philips Research
Image Processing and Network Architecture Dept.
Briarcliff Manor, NY 10510

We describe MAGIC, an experimental system for generating multimedia briefings about the clinical status of post-bypass patients entering a cardiac ICU. MAGIC is a distributed system whose components use knowledge-based techniques for planning and generating briefings in text, speech, and graphics. These briefings are coordinated together by reasoning with dynamically generated temporal and spatial constraints. Formative evaluation using system mock-ups with ICU nurses and residents have been used to determine the general format and content of these briefings. We present an overview of MAGIC's architecture and show what it can currently generate.

INTRODUCTION

Caregivers are tremendous consumers of information, requiring updates on patient clinical status, care plans, and test results. However, it is often difficult for them to obtain only the information that they need in a concise form. Our objective is to develop tools for automatically producing multimedia briefings [1] that meet the information needs of a variety of caregivers, including different specialists and nurses. Through the use of natural-language generation [2], knowledge-based graphics generation [3], and knowledge representation and reasoning systems [4], we are developing an experimental system that can dynamically determine at run time what information to include in a multimedia briefing, how to divide this information among different media, and the form of the language and graphics used.

In the following sections, we first provide an overview of the domain in which we are working and an informal formative evaluation we carried out with potential users of our system. We then describe the software architecture and close with a case study of what the system can currently generate.

DOMAIN

MAGIC (Multimedia Abstract Generation for Intensive Care) was designed to provide post-operative information to hospital caregivers about CABG (Coronary Artery Bypass Graft) patients. Just prior to the completion of CABG surgery and before the patient is brought to the Cardiac ICU (Intensive Care Unit), a number of medical providers must be updated on the patient's status. These include ICU nurses, who must prepare for the patient's arrival, and the patient's cardiologist, who is offsite during surgery and must be updated and informed about the patient's condition prior to the formulation of post-operative clinical management plans. Other administrators and care specialists, who must assure that the appropriate personnel and supplies are available in the ICU before the patient arrives, also need to be informed.

During the critical hour just prior to patient transport, only those caregivers who were present at the operation can provide this information. However, they are typically busy and often unable to do so. This is the time period when the patients are the most tenuous. The last hour in the operating room and the first hour in the ICU are critical to the patient's recovery. The patient scenario is similar to the clinical constraints of outpatient trauma treatment, where expediting and coordinating patient transport and therapy is paramount. There are clear advantages to providing timely, automated briefings for a variety of caregivers.

Our work takes advantage of the existing information infrastructure in the cardiac OR (operating room) at Columbia Presbyterian Medical Center. Patient vital signs are monitored and automatically entered at regular intervals into the LifeLog database system (Modular Instruments Inc.). This data-acquisition system records operative events, procedures, medications, and quality of care issues. In addition, the patient's pre-operative condition is manually entered by the anesthesiologist. Thus, on exiting the operating room, a detailed record of the patient's status before, during, and on completion of surgery is available on-line. We are interested in determining how to extract a selected set of this data and prepare it for presentation to caregivers in a concise and easily understandable form. Among the available data are vital signs, bolus drugs, post-operative drugs, intravenous lines, information about devices such as a pacemaker or balloon pump, data from echocardiograms, and severity assessments.

We hypothesized that an appropriate combination of written text, speech, static graphics, and animation,

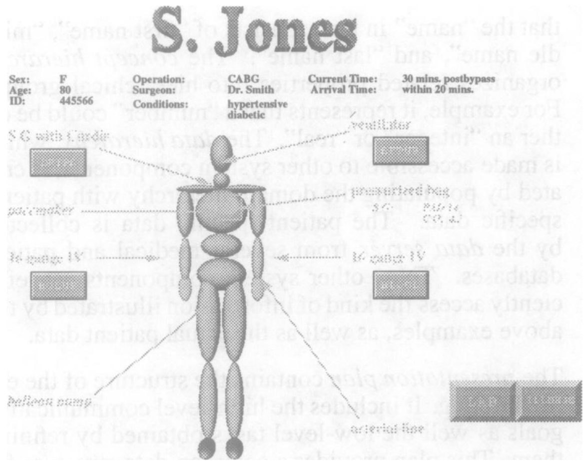

Figure 1: Layout for ICU nurse.

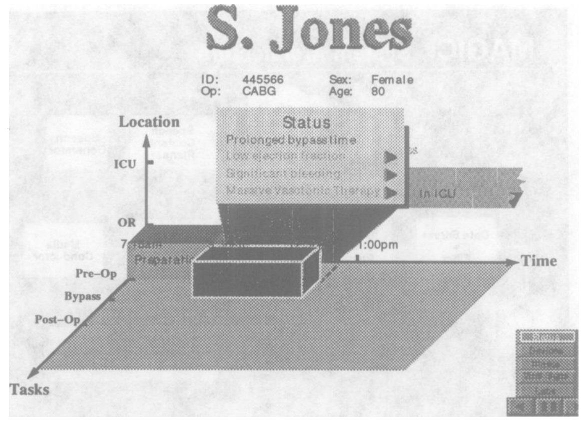

Figure 2: Layout for physician.

coordinated in a coherent multimedia briefing, could be more effective for communicating the desired information than any of these individual media in isolation. Given that different caregivers have different tasks to carry out, one problem in developing a multimedia briefing system is determining which information should be presented to which caregivers and how should it be organized. ICU nurses, for example, must set up to receive the patient in the ICU ward, adjust lines and be ready to immediately initiate or continue post-operative drugs. They are concerned with the exact procedures and medical orders to be carried out to meet the immediate needs of the patient. Physicians, on the other hand, are concerned with more prognostic factors such as length of bypass operation, various vital signs, and severity assessments, all of which indicate how well the patient is doing, patient response to medications, and whether the operation went as expected. Such indicators change the treatment physicians prescribe.

To provide a characterization of the data needed for each caregiver and to determine how important each piece is, we developed an outline based on Weed's problem list format. This SOAP (Subjective, Objective, Assessment, Plan) outline is taught to physicians during medical school, and to other caregivers throughout the medical domain. It requires that subjective information be presented first, then actual data, such as lab reports, followed by assessment, and finally by an actual plan of care. For nurses, a different outline was developed, which specifies that tasks to carry out for the benefit of the patient be presented first, followed by a more cursory objective and assessment report.

Informal Formative Evaluation

To test our hypotheses about what data to present, strategies for order of presentation, and how different media can be used effectively, we carried out an informal formative evaluation with both ICU nurses and physicians (residents) who serve in the cardiac ICU.

For nurses, we presented information using a spatial layout centered around the patient's body, which was used to organize the various lines and devices (Fig. 1). We varied the amount of coordination between speech, text and graphics among the various versions, testing, for example, whether text on the screen should duplicate speech. We used an approach where graphical information was incrementally presented with each successive screen. For physicians, we used an organization of information along a time line (Fig. 2) as well as a more abstract representation of patient status (not shown). The time-line representation showed status and procedures at different points during the course of the operation, using insets and blow-ups to provide details on information such as vital signs or bolus drugs at specific points in the operation.

Physicians indicated a definite preference for the time-line presentation. Nurses indicated that they wanted the summary display first, with all information shown at once in overview form and details provided in successive screens. In addition, we were told that we needed to provide information as concisely as possible in speech, since it takes time to hear. Their comments indicated that precise, full descriptions of the particular devices (e.g., "ventricular pacemaker") could be shown in text while speech could use shorter, more colloquial expressions (e.g., "pacemaker").

SYSTEM ARCHITECTURE

MAGIC consists of several distributed system components that work cooperatively on different sub-tasks of generating multimedia briefings (Fig. 3). The entire system is integrated using ILU (Inter-Language Unification package) [5], which makes it possible for the different components to share data structures and communicate efficiently across different hardware platforms and implementation languages. Prototypes of individual components have been implemented and partially integrated.

The *data server* accesses several medical databases for collecting information about the patient's condition.

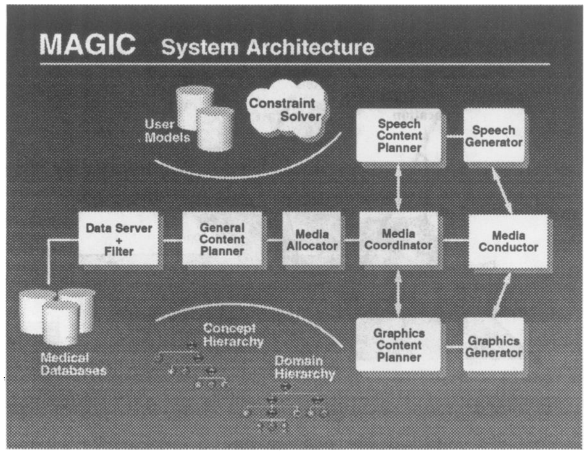

Figure 3: MAGIC system architecture.

The *data filter* selects relevant and important parts of this data, infers some new information, and creates a hierarchical data structure called the *patient data hierarchy* to represent them. The *general content planner* uses the data hierarchy to build a *presentation plan* that expresses the high-level communicative goals to be accomplished by the multimedia briefing. The *media allocator* chooses one or more media to express each communicative goal in the presentation plan.

Each of a set of *medium-specific content planners* can expand the presentation plan by building detailed plans for the communicative goals assigned to its medium. Each medium-specific *media generator* uses these detailed plans to generate its portion of the briefing. The *media coordinator* ensures that the planners and generators of different media are consistent with each other. Finally, the *media conductor* directs the presentation of the material created for each medium so that a single coherent multimedia briefing results.

The databases accessed by the data server include *LifeLog* and patient databases at CPMC. The data filter uses several built-in patient-independent hierarchies, including the *domain hierarchy* and the *concept hierarchy*, to create the patient-specific *data hierarchy*, which is used by all subsequent components. The general content planner selects a plan schema from the *plan library* and instantiates it with the current data to create the presentation plan. Ultimately, various system components will rely on the information in the *user model* to tailor the briefing being developed to the preferences of the current user. They will also rely on a *constraint solver* that represents the constraints among the different parts of the presentation plan and ensures that they are mutually consistent.

The *domain hierarchy*, which is partially drawn from MED (Medical Entities Dictionary) [6], describes the hierarchical grouping of relevant medical and patient data. For example, it represents that the "demographic" information about a patient consists of "name", "medical record number", "gender", and "birth date", and that the "name" in turn consists of "first name", "middle name", and "last name". The *concept hierarchy* organizes related properties into hierarchical groups. For example, it represents that a "number" could be either an "integer" or "real". The *data hierarchy*, which is made accessible to other system components, is created by populating the domain hierarchy with patient-specific data. The patient-specific data is collected by the *data server* from several medical and patient databases. Thus, other system components can efficiently access the kind of information illustrated by the above examples, as well as the actual patient data.

The *presentation plan* contains the structure of the entire briefing. It includes the high-level communicative goals as well the low-level tasks obtained by refining them. This plan provides a common data structure for different components to share information.

The presentation plan includes several temporal and spatial constraints [7]; for example, the demographics information should be displayed before the test results. Although these constraints are generated by several system components and at several different times, a coherent briefing can be generated at the end only if all these constraints are mutually consistent. The *constraint solver* provides the central facility for representing these constraints and determining whether they are consistent. Whenever a constraint is generated in any system component, it is conveyed to the constraint solver, which determines the consistency and returns the result back. The system component may decide to backtrack in case of inconsistency.

Since determining inconsistency is intractable in general, the constraint solver uses an anytime algorithm [4] that sometimes returns the result "maybe consistent". The component that generated the constraint may decide to wait for a definite answer, or may proceed further making some tentative assumption of its own, which it must be willing to revoke when a different definite answer is obtained. Given sufficient time, the constraint solver will always return a definite answer. The other system components can also query the constraint solver to find out the constraints that already exist.

The *media allocator* uses knowledge of various media to distribute high level communicative goals among the media. For example, the goal of conveying several blood pressure values measured during the last 24 hours is assigned to graphics, since some form of a graph or a chart is most suitable for this task. Some important information, for example, the name of the patient is conveyed by both graphics and speech.

The *media coordinator* ensures that the various media-specific content planners and generators are synchronized together. It allows a component to access partial plans developed by other components, so that it can use this information in making its own choices. When an inconsistency among the plans of different components is detected, the media coordinator can make the decision to roll back (backtrack) the plans of some

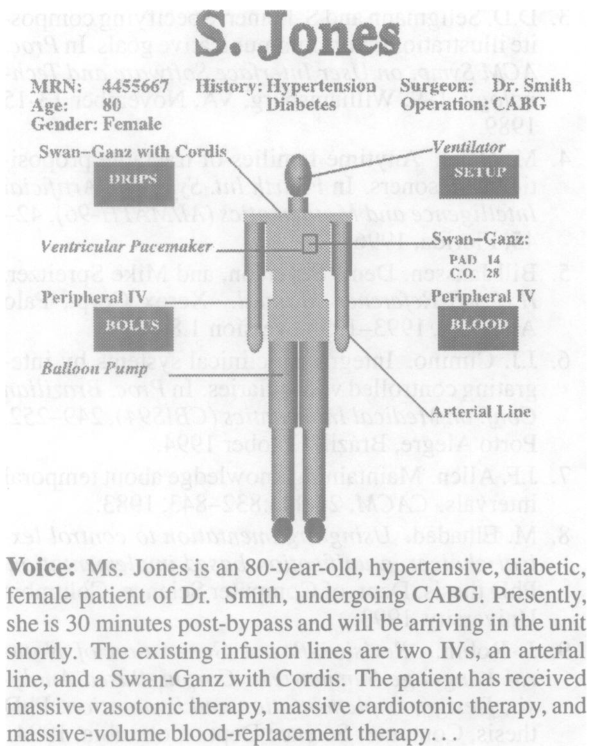

Figure 4: Multimedia presentation generated by MAGIC.

components.

We are planning to add *user models* that will represent preferences of individual users and groups of users about the content and format of the multimedia briefings. In addition, a *media conductor* will "play" the entire presentation plan so that all media are coordinated together.

CASE STUDY

Fig. 4 shows the screen layout as the explanation begins. The text shown at the bottom of the figure is spoken, and the graphics are modified in conjunction with the spoken output. For example, as the demographic information is conveyed, the entry or column in the top table that corresponds to the speech is highlighted as the spoken reference is made. While the first sentence is spoken, the first two columns of the demographics chart are highlighted.

The first step in generating this explanation is to produce a presentation plan using the high-level content planner. MAGIC generates a hierarchical plan containing two main discourse goals: *inform-overview* and *inform-detailview*. These goals are further broken down into substeps that direct the system to provide an overview of demographics, of the time frame, and of the infusion lines. While these overall substeps are temporally ordered by the content planner, the information to be communicated within these steps is unordered and any of the media generators may decide to reorder information within these steps. Note that the presentation plan provides the discourse *segments* for speech. (Currently MAGIC uses stored presentation plans rather than generating them from first principles; in the future, we will investigate how context or information about the user may result in modification of the plan.)

The graphics content planner determines that the demographics are to be presented as a table (currently using a pre-planned format) and the information about the infusion lines and devices are to be presented in the context of an abstract representation of the patient's body. The graphics generator then lays out the table and the representation of the patient.

The goal of both the speech content planner and the speech generator is to produce a concise spoken summary. It is the job of the speech content planner to group information from the plan into sentence-sized chunks. It determines how much information can fit into a sentence and where sentence boundaries should be. To produce a concise summary (one that takes less time to speak), the speech content planner attempts to place more information into a single sentence using modifiers instead of generating separate sentences for each piece of information (which would also be possible). In the first sentence describing demographics, MAGIC groups together six attributes from the LifeLog database (age, gender, medical history, surgeon, procedure, and name) into a single sentence using adjectives to realize attributes where possible. Note that these six attributes all appear under the goal *inform-demographics* in the presentation plan. Thus, medical history appears as two adjectives modifying "female patient" although it could just as easily have been realized as a separate sentence such as, "She has both hypertension and diabetes."

The speech generator has the task of selecting the words and sentence structure to be used. In making word choices, it takes into account that textual references to each object will be generated as well and used as labels on the illustrated object in the accompanying graphics. Thus, for example, speech just says "Two IVs" while the textual labels spell out that they are "peripheral." MAGIC uses Columbia's FUF/SURGE language generation tools for this task [8, 9, 10].

RELATED WORK

A key feature of MAGIC is that both content and form of the presentation are dynamically generated at run-time, thus allowing output to be customized for the current user and situation. In contrast, many interfaces for healthcare applications rely on the more traditional approach of retrieving pieces of canned text or stored images. A few exceptions include research in language generation on tailoring textual descriptions for individual patients. In the HealthDoc project [11], the aim is to produce customized patient education material. Carenini et al. [12] also work on patient specific explanations, having developed a system which can produce natural language descriptions of migraines and respond to follow-up questions requesting further information. The interaction is tailored to the class of migraine patients, the individual patient, and the previous dialogue.

There are a number of systems that, like MAGIC, dynamically generate the content and form of multimedia presentations, but not for healthcare applications. SAGE, which provides textual and graphical presentations of quantitative data, [13, 14], COMET, which provides explanations of equipment maintenance and repair, [15], and WIP, which provides explanation for equipment operation, [16] are all knowledge-based multimedia generation systems that coordinate written text with static graphics. In contrast, in MAGIC, the focus is on coordination of speech and animated graphics, both temporal media. Weitzman and Wittenburg [17] also handle media coordination in their work on generating multimedia presentations. While they can accommodate dynamic relationships among presentation elements, they do not support the ability to replan a presentation as we do in MAGIC, and output is also limited to static text and graphics.

CONCLUSION

We have described MAGIC, a system for generating multimedia briefings describing the status of post-bypass patients entering a cardiac ICU. Preliminary versions of most system components have been implemented and can generate the text and graphics such as that shown in Fig. 4. We are currently integrating the entire system using ILU. We plan to field MAGIC at Columbia-Presbyterian Medical Center and perform user studies for evaluating its effectiveness.

Acknowledgements

MAGIC is an ongoing group project that has benefitted from the design and development work of Shimei Pan, Michelle Zhou, Tobias Höllerer, James Shaw, Yong Feng, and Jeanne Fromer. This research is supported in part by DARPA Contract DAAL01-94-K-0119, the Columbia University Center for Advanced Technology in High Performance Computing and Communications in Healthcare (funded by the New York State Science and Technology Foundation), ONR Contract N00014-94-1-0564, and NSF Grants IRI-94-10117 and GER-90-2406.

References

1. S. Feiner and K. McKeown. Generating coordinated multimedia explanations. In *Proc. IEEE Conf. on AI Applications*, Santa Barbara, CA, March 1990.
2. K.R. McKeown. *Text Generation: Using Discourse Strategies and Focus Constraints to Generate Natural Language Text.* Cambridge Press, Cambridge, England, 1985.
3. D.D. Seligmann and S. Feiner. Specifying composite illustrations with communicative goals. In *Proc. ACM Symp. on User Interface Software and Technology*, 1–9, Williamsburg, VA, November 13-15 1989.
4. M. Dalal. Anytime families of tractable propositional reasoners. In *Fourth Int. Symp. on Artificial Intelligence and Mathematics (AI/MATH-96)*, 42–45, Florida, 1996.
5. Bill Janssen, Denis Severson, and Mike Spreitzer. *ILU 1.8 Reference Manual.* Xerox Corp., Palo Alto, CA, 1993–1995. Version 1.8.
6. J.J. Cimino. Integrating clinical systems by integrating controlled vocabularies. In *Proc. Brazilian Conf. on Medical Informatics (CBIS94)*, 249–252, Porto Alegre, Brazil, October 1994.
7. J.F. Allen. Maintaining knowledge about temporal intervals. *CACM*, 26(11):832–843, 1983.
8. M. Elhadad. *Using argumentation to control lexical choice: a unification-based implementation.* PhD thesis, Dept. of Computer Science, Columbia University, 1993.
9. J. Robin. *Revision-Based Generation of Natural Language Summaries: Corpus-Based Analysis, Design, Implementation and Evaluation.* PhD thesis, Computer Science Dept., Columbia, 1994.
10. K. R. McKeown, M. Elhadad, Y. Fukumoto, J.G. Lim, C. Lombardi, J. Robin, and F.A. Smadja. Text generation in comet. In R. Dale, C.S. Mellish, and M. Zock, eds., *Current Research in Natural Language Generation*, 103–140. Academic Press, 1990.
11. G. Hirst, C. DiMarco, and E.H. Hovy. Authoring and generating heath-education documents that are customized to the individual patient. In *Proc. ITCH '96 Conf.*, Vancouver, B.C., 1996.
12. G. Carenini, V. O. Mittal, and J. D. Moore. Generating patient-specific interactive natural-language explanations. *Journal of the AMIA*, **5**(9), 1994.
13. S.F. Roth, J.A. Mattis, and X. Mesnard. Graphics and natural language as components of automatic explanation. In J.W. Sullivan and S.W. Tyler, editors, *Intelligent User Interfaces*, 207–239. Addison-Wesley, Reading, MA, 1991.
14. V.O. Mittal, S. Roth, J. D. Moore, J. Mattis, and G. Carenini. Generating explanatory captions for information graphics. In *Proc. IJCAI-95*, 1276–1283, Montreal, 1995. Morgan Kaufmann.
15. S. Feiner and K. McKeown. Automating the generation of coordinated multimedia explanations. *IEEE Computer*, 24(10):33–41, October 1991.
16. E. Andre, W. Finkler, W. Graf, T. Rist, A. Schauder, and W. Wahlster. WIP: The automatic synthesis of multimodal presentations. In M. Maybury, ed., *Intelligent Multimedia Interfaces*, 75–93, AAAI/MIT Press, Menlo Park, CA, 1993.
17. L. Weitzman and K. Wittenburg. Automatic presentation of multimedia documents using relational grammars. In *Proc. ACM Multimedia '94*, 403–412, San Francisco, CA, October 1994.